

## A Real-Time Detection Algorithm for Identifying Shill Bidders in Multiple Online Auctions

Nazia Majadi  
Griffith University, Australia  
[nazia.majadi@griffithuni.edu.au](mailto:nazia.majadi@griffithuni.edu.au)

Jarrod Trevathan  
Griffith University, Australia  
[j.trevathan@griffith.edu.au](mailto:j.trevathan@griffith.edu.au)

### Abstract

*Online auctions are highly susceptible to fraud. Shill bidding is where a seller introduces fake bids into an auction to drive up the final price. If the shill bidders are not detected in run-time, innocent bidders will have already been cheated by the time the auction ends. Therefore, it is necessary to detect shill bidders in real-time and take appropriate actions according to the fraud activities. This paper presents a real-time shill bidding detection algorithm to identify the presence of shill bidding in multiple online auctions. The algorithm provides each bidder a Live Shill Score (LSS) indicating the likelihood of their potential involvement in price inflating behavior. The LSS is calculated based on the bidding patterns over a live auction and past bidding history. We have tested our algorithm on data obtained from a series of realistic simulated auctions and also commercial online auctions. Experimental results show that the real-time detection algorithm is able to prune the search space required to detect which bidders are likely to be potential shill bidders.*

### 1. Introduction

An online auction provides ease, comfort and a convenient trading environment to its users. Therefore, online auctions are extremely popular for sellers and buyers. eBay and Yahoo! Auctions are amongst the most popular online auction sites [10]. In the first quarter of 2017, ebay alone has 169 million active users after exceeding 167 million active users by the end of 2016 [21].

However, despite the overwhelming benefits of online auctions, they are attractive to fraudsters. Online auction fraud is one of the fastest growing forms of Internet-based crime. The U.S. Federal Bureau of Investigation's Internet Crime Complaint Center (IC3) reveals that approximately 42 thousand auction-related fraud complaints were received from June 1, 2014 to

December 31, 2014 [22]. Moreover, the latest report of IC3 has found that auction related complaints are still ranking at the top of the complaints list with approximately 116,292 complaints in 2015 [23]. As participants are anonymous, both sellers and bidders can be involved in fraudulent activities. One of the most common types of online auction fraud is shill bidding.

Shill bidding is the act of introducing false bids into an auction on behalf of the seller to artificially raise the item's price so that legitimate bidders must pay more in order to win [14, 15]. The presence of shill bidding also diminishes the reputation of the auction houses as bidders will be reluctant to participate if they feel there is the possibility of being ripped-off by a seller. Shill bidding is difficult to detect because: (i) any user can register under false identity; and (ii) multiple users can form a collusive bidding group to evade detection. Furthermore, it is not easy to prove that someone is indeed guilty of shill bidding.

Shill bidding detection and prevention mechanisms have been proposed by researchers [1, 2, 4, 5, 7, 8, 11, 15-18]. There is now common consensus of the common strategies shill bidders behave in. However, most of these efforts focus on examining bidding patterns once an auction has terminated – as such an innocent bidder has already incurred loss as a result of the shill bidding. More recently, proposals have been made to detect shill bidding in real-time [3, 10, 12, 13, 20]. The idea being that actions can be taken while an auction is currently underway in order to warn/deter shill bidding activity, or prevent the auction from completing if serious price inflating behavior is being detected.

This paper proposes a real-time detection algorithm for shill bidders in multiple auctions. The algorithm splits a live auction into four time periods and examines a bidder's bidding behaviors during each period. A *Live Shill Score (LSS)* is calculated for each bidder during each of these periods based on the bidding behaviors of the bidder in the live auction and his/her past bid history. The auctioneer then notifies legitimate bidders and/or takes necessary actions

against suspected shill bidders based on the *LSS*. We have implemented the algorithm and undertaken preliminary tests on simulated data and commercial auction data. Experimental results show that the algorithm can detect potential shill bidders in multiple auctions and take necessary actions to them in run-time.

This paper is organized as follows: Section 2 gives a background on real-time shill detection approaches and the problem motivation for our research; Section 3 presents our proposed *LSS* algorithm for real-time detection of shill bidding in multiple auctions; and also discusses the calculation procedure of *LSS* for each bidder during a live auction. Section 4 shows our experimental setup and preliminary results. Finally Section 5 provides concluding remarks and avenues for future work.

## 2. Background and Problem Motivation

This section describes some existing shill bidding detection techniques and the motivation of our research on combating shill bidding fraud in online auctions.

### 2.1. Related Work

Commercial online auctions (such as eBay) claim to monitor their auctions for shill bidding. However, they do not disclose how their shill detection methods operate, nor do they publicly acknowledge when shill bidding has been occurring in their auctions. This may be due to fear of lost business should people learn the true extent of shill bidding and how much it is costing legitimate bidders. Furthermore, some unscrupulous auctioneers might actually benefit from inflated prices where there is a commission payable based on the final price.

There are two distinct approaches in scientific literature on shill bidding detection: (i) **Offline (or static) method** which runs after an auction has terminated, or; (ii) **Real-time (or dynamic) method** which operates while an auction is in-progress.

Among the offline approaches, Trevathan and Read [15] present an algorithm to determine the degree of shill bidding in online auctions. The algorithm observes the bidding patterns of each bidder over a series of auctions held by a specific seller and calculates a *Shill Score* for each bidder [15]. The Shill Score is used to indicate the likelihood of the user being a potential shill bidder [15]. Bidders can observe other bidders' Shill Scores and decide whether they want to participate in the particular seller's auction [15]. However, the algorithm works only when an auction ends.

Xu et al. [20] analyse some shill bidding patterns and introduce a *Dynamic Auction Model* (DAM) to detect shill bidding in real-time. Their proposed approach employs three sources: (i) The auction model which is updated dynamically when new bids arrive; (ii) *Linear Temporal Logic* (LTL) formulas represent the shill bidding patterns; and (iii) A *Simple Promela Interpreter* (SPIN) model checker that checks whether the LTL formulas are violated or not [20]. Nevertheless, the problem of this approach is that monitoring an auction after every single submitted bid may make the detection process computationally intensive, which is not practical when there are a large number of participants.

In contrast, Ford et al. [3] present a real-time classifier that identifies suspicious shill bidders based on a neural network method. The neural network is initialized with a training sample and is then updated incrementally to adapt to new bidding data in real-time [3]. However, manually labelling the large training dataset of users is inefficient.

Mamun and Sadaoui [10] propose software architecture to protect auction systems from shill bidders in online auctions. They claim that their proposed mechanism keeps the auction system secure from auction fraud and also maintains trust among users and the online auction system [10]. However, the paper continually mentions an IP tracker that tracks the IP addresses of sellers and bidders. However, the authors do not actually discuss how to implement IP tracker for testing the proposed system. Furthermore, they do not test the effectiveness of their proposal.

Sadaoui et al. [12] propose a generic framework that covers real-time monitoring of multiple live auctions. This framework observes the running auctions, takes actions in real-time and prevents shill bidders from succeeding [12]. The authors state that real-time monitoring processes a smaller number of bids rather than offline processing [12]. This suggests their proposed system can react quicker to the potential shill bidders. However, their approach is a combination of the offline approach with an online approach. Therefore, claims about the speed of their proposal are contradictory. Moreover, they do not address collusion among bidders and/or sellers.

Later Sadaoui et al. [13] propose runtime stage-based monitoring system that detects in-auction fraud by monitoring each bidder's stage activities in ongoing auctions. It then takes immediate action by warning dishonest bidders and cancelling the suspected auctions after detecting abnormal activities in ongoing auctions [13]. However, the authors do not justify the reason for selecting eight bidding behaviors among 17 proposed suspicious bidding patterns.

## 2.2. Problem Motivation

Considering the limitations of the existing literature on shill bidding detection, it is necessary to revise the existing shill detection methods to detect shill bidding before any payment. This approach will potentially prevent monetary loss for victims. Moreover, it is preferable to detect and stop shill bidders from undertaking shill bidding in the run-time rather than detect it afterwards.

The research questions we are seeking to address are as follows:

- (i) How will the proposed shill bidding detection method be able to deter shill bidders in real-time?
- (ii) What are the typical shill bidding patterns that can be used in determining whether shill bidding is happening in real-time?
- (iii) What are the consequences for taking disciplinary actions against a shill bidder after detecting his/her involvement in a particular live auction?

## 3. Methodology

Our shill detection methodology considers bidding properties while an auction is currently in progress. We propose a *Live Shill Score (LSS)* that represents a bidder's bidding behavior during a live auction. The *LSS* is calculated for each bidder based upon a selection of bidding behaviors/characteristics and the bidder's past bid history. The *LSS* can be used to potentially identify shill bidders, and enforce disciplinary action during an auction.

### 3.1. Assumptions

To develop our *LSS* algorithm, we make the following assumptions:

- (i) There is one shill bidder;
- (ii) There are multiple auctions (one live auction and rest of them are past auctions); and
- (iii) There is one seller (note that the seller is not the same entity as the auctioneer).

Our proposed algorithm is not addressing situations where there are concurrent auctions running for the same item (as in [19]).

### 3.2. Monitoring Stages of a Live Auction

A shill bidder submits bids at certain times throughout the auction which can be deliberated more conducive to price inflating behavior. For instance, it

becomes more unsafe for shill bidders to submit fake bids closer to the end of an auction as there is an increased possibility that the shill bidder will win the auction if he/she is not outbid in time. Therefore, our proposed *LSS* algorithm splits an auction into a series of stages depending on the time elapsed for reacting to a shill bidder and other legitimate bidders in real-time. This strategy was first suggested in Xu et al.'s real-time shill detection proposal [20]. The different auction stages are as follows:

- (i) **Early Stage** – This is the first quarter of the auction duration (up to 25% of the auction duration). A shill bidder usually submits bids early in an auction to encourage legitimate bidders to enter higher bids in order to win.
- (ii) **Middle Stage** – This is the time period between 25.1% and 80% of the auction duration. A shill bidder places most of his/her bids in this stage.
- (iii) **Late Stage** – The next 15% (between 80.1% and 95%) of the auction is the late stage. A shill bidder tries to avoid placing bids during this period to reduce his/her chance of winning the auction.

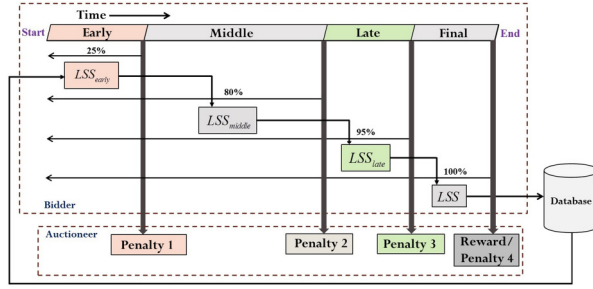
In our real-time shill bidding detection methodology, we consider the above stages for monitoring shill bidding patterns in a live auction. However, we differ from Xu et. al [19, 20] by introducing an additional stage, referred to as the *Final stage*:

- (iv) **Final Stage** – The last 5% of the auction duration is used for verifying the detected bidding patterns from the early, middle and late stages. The overall *LSS* of each bidder in the auction will be calculated during this stage.

The reason for the final stage is that Xu et. al [20] did not clearly address when reactions would be taken against the potential shill bidders after the end of the late stage. Therefore, the final stage is required for taking appropriate actions against potential shill bidders after the late stage and for exonerating bidders with good behavior.

### 3.3. The Real-Time Shill Bidding Detection Process

Figure 1 illustrates the process for real-time shill bidding detection in multiple auctions. Time flows from left to right.



**Figure 1: Real-time shill bidding detection process in multiple auctions.**

A live auction is split into the four aforementioned stages (i.e., early, middle, late, final). Each bidder's behavior is examined against other bids which are placed by the bidder during each auction stage. We also consider the bidder's past bidding history for calculating each bidder's  $LSS$ . An  $LSS$  for each bidder is calculated during each respective auction stage. The  $LSS$  for each bidder in a particular auction stage is as follows:

- (i)  $LSS_{early}$  – After completing 25% of auction duration,  $LSS_{early}$  is calculated based on the bidding behavior of each bidder during the early stage of the live auction and past bidding history of the bidder.
- (ii)  $LSS_{middle}$  – The value of  $LSS_{middle}$  is calculated after completing the middle stage of the auction duration. The calculation of  $LSS_{middle}$  is based on the bidding behavior of each bidder during the middle stage of the live auction and past bidding history of the bidder.
- (iii)  $LSS_{late}$  –  $LSS_{late}$  is calculated after completing the late stage. The calculation of  $LSS_{late}$  is based on the bidding behavior of each bidder during the late stage of the live auction and past bidding history of the bidder.
- (iv)  $LSS$  –  $LSS$  is calculated based on the bidding behaviors of all bidders throughout the entire auction and the bidding behavior of the bidders in past auctions.

A high value of a bidder's  $LSS$  indicates the increased likelihood the bidder is engaging in price inflating behavior. At the end of the auction, a bidder's overall  $LSS$  is stored into the database.

### 3.4. Reactions to shill bidders

After every stage of the auction, the algorithm will determine whether any action is required based on the bidding behavior so far in the auction. Apart from Xu

et al. [19, 20], to our knowledge there is no literature discussing potential penalties or the impact of these penalties on bidders' bidding behavior. We propose the following penalties for our algorithm:

- (i) **Penalty 1 (At the end of the early stage)** – Send a warning message to the suspected bidders. This penalty encourages honest bidders to continue bidding without causing anxiety. A potential shill bidder would also realize that the auction is being monitored, thereby forcing him/her to rethink his/her behavior.
- (ii) **Penalty 2 (At the end of the middle stage)** – Postpone the auction temporarily; notify all participating bidders about the suspicious activity; and also ask suspicious bidders and/or the seller to show cause for why the auction should continue. This penalty has a stronger psychological impact on auction participants. The penalty sends a message that the auction is potentially at risk of shill bidding. Therefore, everyone is on notice to improve their behavior. At this point, honest bidders can decide whether they want to remain involved with the auction or not. On the other hand, the suspicious bidders and/or seller would realize that they need to improve their behavior. Otherwise the postponed auction will not be reopened.
- (iii) **Penalty 3 (At the end of the late stage)** – Stop the current auction permanently. This penalty also limits the privileges of the suspicious account when couple of offences occurred. For example, feedback rating of a shill bidder can be reduced as it is an important part of eBay community to evaluate a seller's reputation. This penalty informs honest bidders this is a monitored and safe bidding system. The suspicious bidders and/or seller would realize that they need not only to show their good behavior to run auctions in future but also to get back good reputation. Otherwise he/she will not be able to list any auction further. The penalty reduces the potential for monetary loss.
- (iv) **Reward/Penalty 4 (At the end of the final stage)** – Suspend the suspected account permanently if it shows repeated behaviors in multiple live auctions. In addition, bidders whose behavior appears to be normal are exonerated (i.e., the winner of the auction, late bidders, bid snipers).

In addition, there can be some economic penalties for the potential shill bidders.

### 3.5. Shill bidding behaviors

There are many shill bidding strategies found in past studies. We identified the seven most popular shill bidding characteristics based on an analysis of the literature [9].

Table 1 shows the bidding behaviors with the corresponding ratings considered for our *LSS* algorithm. Among the seven bidding patterns [9], ‘early bidding’ and ‘last/late bidding’ indicate the same bidding characteristic. Therefore, we consider ‘early bidding’ in our *LSS* algorithm.

**Table 1: Bidding behavior for real-time detection of shill bidding**

Rating	Bidding Behaviors
$\alpha$	Affinity to a seller
$\beta$	Bid frequency
$\gamma$	Win/lose factor
$\delta$	Rapid outbid time
$\varepsilon$	Small bid increment
$\zeta$	Early bidding

**Table 2: Bidding behaviors at each auction stage**

Stages	Bidding Behaviors
Early Stage	Affinity to a seller ( $\alpha$ rating)
Middle Stage	Bid frequency ( $\beta$ rating)
Late Stage	Rapid outbid time ( $\delta$ rating)
	Small bid increment ( $\varepsilon$ rating)
	Early bidding ( $\zeta$ rating)
Final Stage	Affinity to a seller ( $\alpha$ rating)
	Bid frequency ( $\beta$ rating)
	Win/Lose factor ( $\gamma$ rating)
	Rapid outbid time ( $\delta$ rating)
	Small bid increment ( $\varepsilon$ rating)
	Early bidding ( $\zeta$ rating)

Table 2 shows the bidding behaviors considered at different stages (i.e., early, middle, late and final stage) for calculating the *LSS* of each bidder in a live auction. The following bidding characteristics are considered for calculating  $LSS_{early}$ ,  $LSS_{middle}$ , and  $LSS_{late}$  for each bidder in a live auction:

- (i) **Affinity to a seller ( $\alpha$  rating)** – A shill bidder usually participates in auctions run by one particular seller. A bidder with a high  $\alpha$  rating is suspicious to be a shill bidder. Suppose  $m$  auctions held by a seller. To calculate  $\alpha$  rating, we count the number of the seller’s auctions,  $m^i$ , bidder  $i$  has participated in. Then,  $\alpha$  rating of bidder  $i$  can be calculated as follows:

$$\alpha^i = (m^i - w^i)/m; \quad 0 \leq \alpha^i \leq 1$$

where  $m^i$  represents the number of auctions bidder  $i$  participated in run by a seller and  $w^i$  denotes the number of auctions bidder  $i$  won. In general,  $\alpha$  rating will be high for a shill bidder.

- (ii) **Bid frequency ( $\beta$  rating)** –  $\beta$  rating indicates the average percentage of bids that a bidder has submitted throughout a particular stage of the live auction and over the past auctions he/she had participated in.

For calculating the  $\beta$  rating, we count the number of auctions,  $m^i$ , a bidder  $i$  has participated in. Suppose  $M$  is the set of auction numbers. The  $\beta$  rating of bidder  $i$  across all auctions,  $j \in M^i$ , can be calculated as follows:

$$\beta^i = \begin{cases} 0, & m^i = 0 \\ \frac{1}{m^i} \sum_{j \in M^i} \frac{n_j^i}{n_j}, & \text{otherwise} \end{cases}; \quad 0 \leq \beta^i \leq 1$$

where  $n_j^i$  denotes the number of bids placed by bidder  $i$  in auction  $j$  and  $n_j$  is the total number of bids placed by all bidders in auction  $j$ . In general, the  $\beta$  rating will be high for a shill bidder compared to a legitimate bidder.

- (iii) **Rapid outbid time ( $\delta$  rating)** – A shill bidder gives more time to legitimate bidder for responding on his/her fake bids. Therefore, a shill bidder submits a new bid within a small time period of a genuine bid. This behavior denotes  $\delta$  rating which can be measured by observing inter bid times for all bidders. The average inter bid time is found for each bidder across a particular stage of the live auction and the past auctions he/she had participated in. Bidders who wait longer between bids have a lower average inter bid time score.

For calculating the  $\delta$  rating, we count the number of auctions,  $m^i$ , a bidder  $i$  has participated in. Suppose  $M$  is the set of auction numbers. The  $\delta$  rating of bidder  $i$  across all auctions,  $j \in M^i$ , can be calculated as follows:

- (a) We calculate the inter bidding time for each bid,  $k$ , submitted by bidder  $i$  in auction  $j$ :

$$\Delta t_{j,k}^i = \begin{cases} 0, & k = 1 \\ t_{j,k}^i - t_{j,k-1}^i, & k > 1, k \in N_j \end{cases}$$

where  $t_{j,k-1}^i$  is the time of a previous bid submitted by a rival bidder. Note that,  $N_j$  be the bid numbers (e.g. 1<sup>st</sup> bid, 2<sup>nd</sup> bid, etc) in auction  $j$  and  $|N_j| = n_j$ .

- (b) Then, we calculate the average inter bidding time for bidder  $i$  in auction  $j$ :

$$\delta_j^i = \frac{1}{n_j^i} \sum_{k \in N_j^i} \Delta t_j^i$$

where  $N_j^i$  is the bid numbers of bidder  $i$  in auction  $j$ .

So, the final  $\delta$  rating of any bidder  $i$  over all auctions participated in:

$$\delta^i = 1 - \left( \frac{1}{m^i} \sum_{j \in M^i} \delta_j^i \right); \quad 0 \leq \delta^i \leq 1$$

In general, the  $\delta$  rating will be higher for a shill bidder in comparison to a legitimate bidder.

- (iv) **Small bid increment ( $\varepsilon$  rating)** – A shill bidder outbids a legitimate bid by a large amount will increase the risk of losing an auction [8]. Therefore, a shill bidder tends to only bid the minimal amount to stay ahead of the leading bid. This behavior denotes  $\varepsilon$  rating which can be measured by observing inter bid increments for all bidder. The average inter bid increment is found for each bidder across a particular stage of the live auction and the past auctions he/she had participated in. Bidders who place bids smaller bid increments have a lower average inter bid increment score.

For calculating the  $\varepsilon$  rating, we count the number of auctions,  $m^i$ , a bidder  $i$  has participated in. Suppose  $M$  is the set of auction numbers. The  $\varepsilon$  rating of bidder  $i$  across all auctions,  $j \in M^i$ , can be calculated as follows:

- (a) We calculate the inter bid increments for each bid,  $k$ , submitted by bidder  $i$  in auction  $j$ :

$$\Delta p_{j,k}^i = \begin{cases} 0, & k = 1 \\ p_{j,k}^i - p_{j,k-1}^i, & k > 1, k \in N_j^i \end{cases}$$

where  $p_{j,k-1}^i$  is the price of a previous bid submitted by a rival bidder.

- (b) Then, we calculate the average inter bid increment for bidder  $i$  in auction  $j$ :

$$\varepsilon_j^i = \frac{1}{n_j^i} \sum_{k \in N_j^i} \Delta p_j^i$$

where  $N_j^i$  is the bid numbers of bidder  $i$  in auction  $j$ .

So, the final  $\varepsilon$  rating of any bidder  $i$  over all auctions participated in:

$$\varepsilon^i = 1 - \left( \frac{1}{m^i} \sum_{j \in M^i} \varepsilon_j^i \right); \quad 0 \leq \varepsilon^i \leq 1$$

In general, the  $\varepsilon$  rating will be higher for a shill bidder in comparison to a legitimate bidder.

- (v) **Early bidding ( $\zeta$  rating)** – If a shill bidder bids towards the end of an auction, he/she risks of losing the auction by not being outbid in time. Therefore, a shill bidder tries to bid early in an auction as it is less risky and maximizes the amount of influence the shill has over an auction. The  $\zeta$  rating denotes how early in an auction a bidder commenced bidding.

To calculate the  $\zeta$  rating of a bidder  $i$ , we count the number of auctions,  $m^i$ , the bidder  $i$  has participated in. Suppose  $M$  is the set of auction numbers. The  $\zeta$  rating of bidder  $i$  across all auctions,  $j \in M^i$ , can be calculated as follows:

We calculate the difference between auction  $j$ 's starting time,  $t_j$ , and the time,  $t_{j,k_0}^i$  of the first bid placed by bidder  $i$  in auction  $j$ .

$$\zeta_j^i = t_{j,k_0}^i - t_j$$

Finally, we calculate the average of the final bid time differences for bidder  $i$  over all auctions participated in:

$$\zeta^i = 1 - \frac{1}{m^i} \sum_{j \in M^i} \zeta_j^i; \quad 0 \leq \zeta^i \leq 1$$

In general, the  $\zeta$  rating will be high for a shill bidder.

We consider the following additional bidding behavior for calculating the overall  $LSS$  at the end of a live auction:

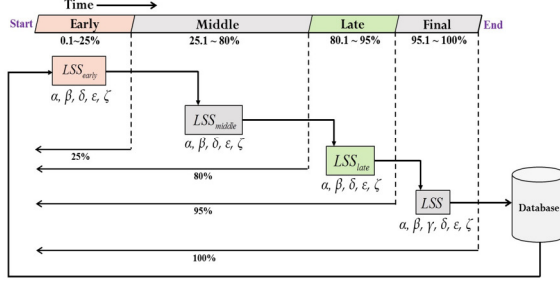
- (vi) **Win/lose factor ( $\gamma$  rating)** – A shill bidder avoids winning an auction, as the auction will have to be repeated. A bidder with a low  $\gamma$  rating indicates suspicious behavior. It can be mentioned that a winner of an auction is not a shill bidder. A bidder  $i$  who has won  $w^i$  auctions, the  $\gamma$  rating of the bidder  $i$  is calculated as follows:

$$\gamma^i = 1 - \left( w^i / m^i \right); \quad 0 \leq \gamma^i \leq 1$$

In general, a shill bidder will have a high  $\gamma$  rating.

### 3.6. Calculation of the $LSS$

After calculating  $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\varepsilon$ ,  $\zeta$ , and  $\gamma$  ratings in a particular stage of auction duration, then we can calculate the value  $LSS$  after completing every stage (see Figure 2).



**Figure 2: LSS calculation process**

In the early stage,  $LSS_{early}$  is calculated as follows:

$$LSS_{early} = \frac{\theta_1\alpha + \theta_2\beta_{early} + \theta_3\delta_{early} + \theta_4\varepsilon_{early} + \theta_5\zeta_{early}}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5} \times 10 \quad (1)$$

At the middle stage,  $LSS_{middle}$  is calculated as follows:

$$LSS_{middle} = \frac{\theta_1\alpha + \theta_2\beta_{middle} + \theta_3\delta_{middle} + \theta_4\varepsilon_{middle} + \theta_5\zeta_{middle}}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5} \times 10 \quad (2)$$

At the late stage,  $LSS_{late}$  is calculated as follows:

$$LSS_{late} = \frac{\theta_1\alpha + \theta_2\beta_{late} + \theta_3\delta_{late} + \theta_4\varepsilon_{late} + \theta_5\zeta_{late}}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5} \times 10 \quad (3)$$

At the final stage,  $LSS$  is calculated as follows:

$$LSS = \frac{\theta_1\alpha + \theta_2\beta + \theta_3\delta + \theta_4\varepsilon + \theta_5\zeta + \theta_6\gamma}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5 + \theta_6} \times 10 \quad (4)$$

where  $\theta$  is the weighted coefficient associated with each rating in Equations (1), (2), (3) and (4). The weights used in this paper are:  $\theta_1 = 9$ ,  $\theta_2 = 2$ ,  $\theta_3 = 2$ ,  $\theta_4 = 2$ ,  $\theta_5 = 2$ , and  $\theta_6 = 5$ . These weightings are obtained by experimenting with simulated auction data and also by looking at historical commercial auction data. Note that it is difficult to calibrate weightings from known auctions that contain shill bidding as such data is extremely limited (due to non-disclosure by commercial auctioneers). We chose the weightings using the following justifications:

A bidder's  $\alpha$  rating receives the highest weighting. This is done to count one-time bidders who participate most of auctions run by a particular seller and always submitted bid in the early and middle stages of an auction.

A bidder's  $\gamma$  rating is given the next highest weighting because it is the strongest sign that the winner of the auction is not a shill bidder.

The next significant factors are the  $\beta$ ,  $\delta$ ,  $\varepsilon$ ,  $\zeta$  ratings. These factors are used to see if the bidding behavior matches a shill bidder's behavior. A bidder's  $\beta$  rating will be high for an aggressive shill bidder or low for one-time bidder. It may happen that a one-time bidder submits bid most of the auctions either early or middle stage of the auction duration. As a result, the  $\beta$  rating can present mixed results depending on the type of shill behavior employed. For this reason, the  $\beta$  rating is given a lower weighting to consider both type of bidders.

The  $\delta$  and  $\varepsilon$  weightings receive lower rating than  $\alpha$  and  $\gamma$  weightings because of the effect of one-time

bidders. That is, if a bidder only bids once at any stages of a live auction, submitting the bid quickly just after the current highest bid and by minimal amount required, then the bidder will have high  $\delta$  and  $\varepsilon$  ratings. As the bidder does not bid again, his/her  $\delta$  and  $\varepsilon$  ratings will always remain high and not average down if the bidder submits slower and larger bids later. Therefore, the weightings of  $\delta$  and  $\varepsilon$  highly results in many one-time bidders scoring high overall even though such bidders are clearly innocent.

A bidder's  $\zeta$  weightings are also given two as it is not more influential than other bidding behavior ratings. Instead all the bidding characteristic weightings must be examined as a group to determine if the bidder's bidding behavior fits the profile of a shill bidder.

The  $LSS$  values in all stages of a live auction are between 0 and 10. The bidder with the highest  $LSS$  is the most likely to be a potential shill bidder.

## 4. Experimental Results

The  $LSS$  algorithm was tested on simulated and commercial auction data. The first test involved a series of simulated auction trials which were acquired from Trevathan and Read [15]. The second test considered commercial online auctions which were collected from Jank and Shmueli [6].

### 4.1. Simulated auctions

Trevathan and Read [15] conducted a series of simulated auction trails to obtain auction data. The data contained thirty-nine auctions where each of these auctions was for a different item. All auctions were considered to be for one seller [15].

The auction proceedings involved twenty-six bidders. A bidder's goal was to win while also trying to save his/her money. On the other hand, a shill bidder's goal was to force a bidder into spending as much of his/her money as possible. Bidders were not informed that shill bidding was occurring. Furthermore, the shill bidder had no knowledge of how much money bidders had.

There were two types of tests performed: (i) with one shill bidder, and; (ii) without shill bidding. The  $LSS$  algorithm was run on all of the tests to determine its effectiveness on determining shill bidder and the likelihood of it incriminating innocent bidders. Each bidder's  $LSS$  was compared to a predetermined threshold value for a known shill bidder. The threshold values were different at four different stages of a live auction. For example,  $\rho_{t_{early}}$  for the early stage,  $\rho_{t_{middle}}$  for the middle stage,  $\rho_{t_{late}}$  for the late stage,



and  $\rho_{t_{final}}$  for the final stage. To experiment  $LSS$  algorithm on simulated auction data, we consider  $\rho_{t_{early}} = 9$ ,  $\rho_{t_{middle}} = 8$ ,  $\rho_{t_{late}} = 8$ , and  $\rho_{t_{final}} = 7$ . These values were obtained by experimenting with simulated and commercial auction data. Earlier auction stages have higher thresholds to ensure that innocent bidders do not incur penalties due to the lack of evidence.

**4.1.1. Auctions with shill bidding.** The first test involved ten auctions and sixteen bidders, including one shill bidder (i.e., ‘Shelly’). Table 3 presents the Shill Scores for each bidder at the end of the auction (as outlined in [15]). Table 3 shows that ‘Shelly’ is identified as engaging in the most significant price inflating behavior (i.e., ‘Shelly’ has a Shill Score of 9.16).

**Table 3: Shill Scores for each bidder using Trevathan and Read approach [15]**

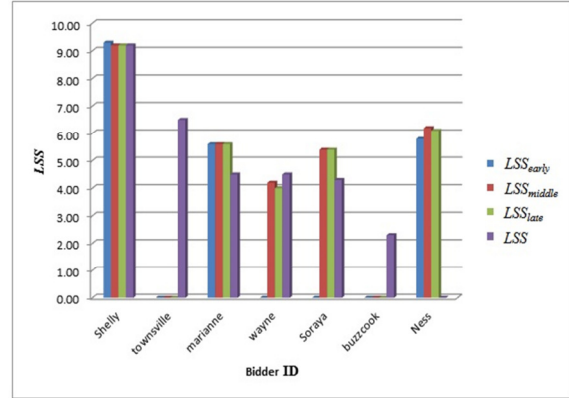
Bidder ID	$\alpha$	$\beta$	$\gamma$	$\delta$	$\varepsilon$	$\zeta$	Shill Score
Shelly	1.00	0.70	0.90	0.96	0.88	0.79	9.16
townsville	0.60	0.21	0.83	0.86	0.65	0.60	6.45
Marie	0.60	0.11	0.83	0.45	0.89	0.02	5.84
jc112425	0.40	0.20	0.75	0.75	0.72	0.79	5.58
Ness	0.40	0.48	0.00	0.88	0.90	0.85	4.46
wayne	0.20	0.42	0.50	0.75	0.93	0.66	4.46
marianne	0.60	0.18	0.00	0.62	0.91	0.39	4.36
Soraya	0.40	0.32	0.00	0.91	0.85	0.78	4.24
Bear	0.20	0.22	0.50	0.98	0.16	0.43	3.58
brenda	0.20	0.59	0.00	0.97	0.97	0.41	3.49
joe	0.10	0.67	0.00	0.62	0.92	1.00	3.33
banana	0.10	0.11	0.00	1.00	1.00	1.00	3.24
ronisanidiot	0.10	0.10	0.00	0.95	0.98	0.76	2.95
hrdw	0.10	0.10	0.00	0.99	0.99	0.68	2.92
buzzcook	0.10	0.08	0.00	1.00	0.99	0.01	2.30
groper	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Although Trevathan and Read’s approach [15] was able to detect a potential shill bidder, their Shill Score algorithm does not work in real-time. Therefore, we chose to run our  $LSS$  algorithm on the same auction data to see how it would execute in detecting a shill bidder in real-time.

To experiment with the  $LSS$  algorithm, we considered one live auction (i.e., *Auction ID- 9*) among the ten auctions and remaining nine auctions were considered as past auctions. Table 4 shows the  $LSS$  of each bidder from *Auction ID- 9*. ‘Shelly’ showed shill bidding behavior in each stage of the auction duration (i.e., consistently high scores (Figure 3)). This results in the  $LSS$  algorithm postponing the auction after the middle stage due to generating *Penalty 2*. ‘Shelly’ is then put on notice before the auction resumes. Therefore, the  $LSS$  takes actions against a potential shill bidder earlier as compared to [15].

**Table 4:  $LSS$  of each bidder in the different stages of Auction ID- 9**

Bidder ID	$\alpha$	$\beta$	$\gamma$	$\delta$	$\varepsilon$	$\zeta$	$LSS_{early}$	$LSS_{middle}$	$LSS_{late}$	$LSS$
Shelly	1.00	0.72	0.90	0.96	0.88	0.79	9.30	9.20	9.20	9.20
townsville	0.60	0.05	0.83	0.22	0.16	0.15	0.00	0.00	0.00	6.50
marianne	0.60	0.25	0.00	0.59	0.91	0.37	5.60	5.60	5.60	4.50
wayne	0.20	0.34	0.50	0.58	0.72	0.35	0.00	4.20	4.00	4.50
Soraya	0.40	0.27	0.00	0.68	0.64	0.54	0.00	5.40	5.40	4.30
buzzcook	0.10	0.02	0.00	0.25	0.25	0.00	0.00	0.00	0.00	2.30
Ness	0.00	0.37	0.00	0.62	0.58	0.58	5.80	6.20	6.10	0.00



**Figure 3:  $LSS$  of each bidder participated in Auction ID- 9**

**4.1.2. Auctions without shill bidding.** The second test involved ten auctions held by a seller. There were 18 bidders. To experiment our  $LSS$  algorithm, we considered one live auction and rest of ten auctions as past auctions. However, unlike the previous test no intentional shill bidding behavior had been engaged in.

Table 5 and Figure 4 show the  $LSS$  of each bidder in *Auction ID- 8*. The test’s purpose is to gauge regular bidding behavior in the different stages of a live auction. Scores are consistently low for all bidders; therefore no real-time intervention is required.

**Table 5:  $LSS$  of each bidder in the different stages of Auction ID- 8**

Bidder ID	$\alpha$	$\beta$	$\gamma$	$\delta$	$\varepsilon$	$\zeta$	$LSS_{early}$	$LSS_{middle}$	$LSS_{late}$	$LSS$
jc112425	0.40	0.42	0.75	0.99	0.90	0.99	6.20	5.90	5.90	6.30
Ness	0.30	0.70	0.67	0.90	0.82	0.92	4.90	5.70	5.70	5.90
brenda	0.30	0.34	0.00	0.64	0.67	0.13	0.00	4.30	4.30	3.50
marianne	0.20	0.26	0.00	0.63	0.64	0.55	0.00	4.20	4.20	3.50
Marie	0.10	0.17	0.00	0.71	0.68	0.05	0.00	3.00	3.00	2.50
buzzcook	0.10	0.12	0.00	0.74	0.65	0.01	0.00	2.90	2.90	2.20
Bear	0.00	0.20	0.00	0.41	0.24	0.25	0.00	4.10	4.10	0.00



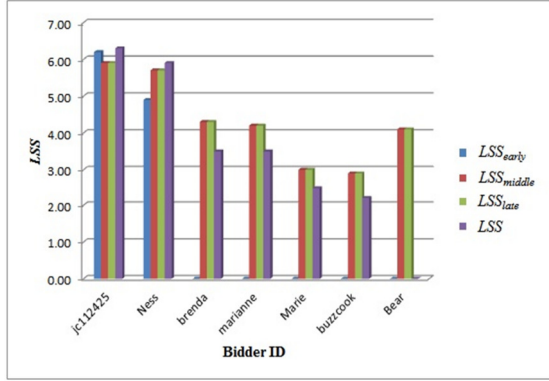


Figure 4: LSS of each bidder in Auction ID- 8

## 4.2. Commercial auctions

The  $LSS$  algorithm was tested on commercial auction data. We considered an auction listing of Palm Pilot PDAs auctioned on eBay [6]. We selected the PDA because of its popularity. The auction data was collected over a period of two months in 2003. The auction data contains 149 auctions across multiple sellers.

We considered one auction as live (i.e., Auction ID- 3023885982 run by ‘michael-33’) among 149 auctions and the rest of them were past auctions. To experiment the  $LSS$  algorithm on commercial auction data, we consider  $\rho_{t_{early}} = 6$ ,  $\rho_{t_{middle}} = 6$ ,  $\rho_{t_{late}} = 6$ , and  $\rho_{t_{final}} = 6$ . These values were obtained by experimenting with commercial auction data.

Table 6 shows the  $LSS$  of each bidder in Auction ID- 3023885982. Figure 5 illustrates that ‘chimam’ showed consistent shill behavior (i.e., scores above 6). The  $LSS$  algorithm triggers Penalties 1 and 2 as a result. Therefore, the  $LSS$  algorithm starts intervening in the auction at the end of the early and middle stages. This intervention could potentially prevent further bad behavior from this bidder.

Table 6: LSS of each bidder in the different stages of Auction ID- 3023885982

Bidder ID	$\alpha$	$\beta$	$\gamma$	$\delta$	$\varepsilon$	$\zeta$	$LSS_{early}$	$LSS_{middle}$	$LSS_{late}$	$LSS$
chimam	0.75	0.08	0.92	0.87	0.45	0.92	6.70	6.70	6.70	7.30
rexth	0.50	0.13	0.87	0.56	0.65	0.77	5.10	5.10	5.10	5.90
andalucas	0.19	0.05	0.67	0.68	0.74	0.67	0.00	4.30	4.40	4.90
charly0	0.25	0.09	0.75	0.70	0.63	0.83	3.80	4.00	4.00	4.80
momassanta	0.25	0.07	0.75	0.67	0.84	0.64	3.80	4.00	4.00	4.80
mikeyokred	0.19	0.25	0.67	0.59	0.68	0.37	0.00	4.00	4.00	4.60
goober-half	0.13	0.09	0.50	0.70	0.62	0.25	0.00	3.20	3.30	3.70
akphotog	0.13	0.03	0.50	0.25	0.19	0.01	0.00	0.00	0.00	3.40
james91722	0.13	0.08	0.50	0.59	0.35	0.08	0.00	2.30	2.40	3.20
jmhboots	0.19	0.11	0.00	0.69	0.72	0.17	0.00	3.60	3.70	2.80
afroitaliano31	0.13	0.04	0.50	0.32	0.17	0.08	0.00	0.00	2.10	2.80
csekhara001	0.06	0.07	0.00	0.49	0.42	0.03	0.00	0.00	2.60	2.10
lthornton	0.00	0.02	0.00	0.23	0.08	0.02	0.00	0.00	2.30	0.00

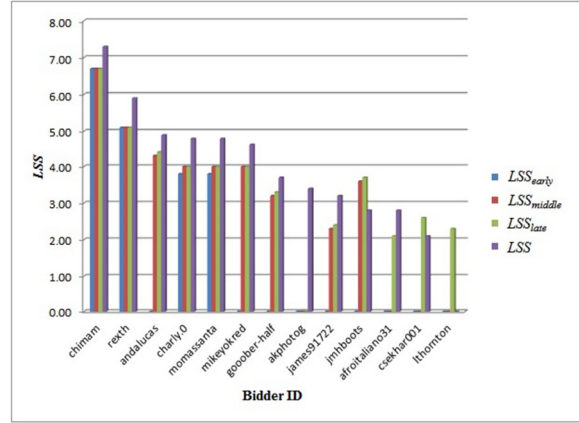


Figure 5: LSS of each bidder in Auction ID- 3023885982

On further inspection, ‘chimam’ has participated in 75% of total auctions run by the same seller (i.e., ‘michael-33’) which is suspicious (see Table 6).

## 5. Conclusion and Future Work

This paper proposed a real-time detection method for combating shill bidding in multiple online auctions. Our  $LSS$  algorithm aims to not only detect potential shill bidders in real-time but also to react to the shill bidding behavior as early as possible during a live auction. Therefore, this approach potentially reduces the monetary loss of shill bidding victims.

We implemented the  $LSS$  algorithm and undertook some experimentation to determine how effective the algorithm is in detecting shill bidders in run-time. The first tests involved simulated auction data (i.e., real users with fake auctions). These tests contained data for known shill bidding and also auctions with no shill bidding. The results show that the  $LSS$  was able to highlight a potential shill bidder during a live auction.

Further testing involved using commercial auction data collected from eBay. It was not known whether this data contained any shill bidders. The results for the  $LSS$  algorithm were consistent with the simulated auction data. The algorithm was able to identify a highly suspicious bidder in run-time.

Future work involves undertaking more comprehensive testing of the  $LSS$  algorithm in multiple auctions run by multiple sellers using a wider range of auction data. Additionally, we would also like to further develop our auction system for detecting and preventing multiple shill bidders from performing price inflating attempts with collusive groups in real-time. Further work also involves looking into the impact of imposing penalties on bidders and extending our proposal to look at concurrent auctions for the same item.

## 6. References

- [1] Dong, F., S.M. Shatz, H. Xu, "Inference of Online Auction Shills using Dempster-Shafer Theory," in Proceedings of the 6<sup>th</sup> International Conference on Information Technology: New Generations (ITNG), 2009, pp. 908–914.
- [2] Dong, F., S.M. Shatz, H. Xu, D. Majumdar, "Price Comparison: A Reliable Approach to Identifying Shill Bidding in Online Auctions?," *Electronic Commerce Research and Applications*, vol. 11, no. 2, 2012, 171–179.
- [3] Ford, B. J., H. Xu, and I. Valova, "A real-time self-adaptive classifier for identifying suspicious bidders in online auctions," *The Computer Journal*, vol. 56, no. 5, 2012, pp. 646–663.
- [4] Ford, B.J., H. Xu, C.K. Bates, S.M. Shatz, "Specification of Flexible and Complex Bidding Strategies in Agent-based Online Auctions," in Proceeding of the 6<sup>th</sup> International Conference on Information Technology: New Generations (ITNG), 2009, pp. 894–900.
- [5] Goel, A., H. Xu, S.M. Shatz, "A Multi-State Bayesian Network for Shill Verification in Online Auctions," in Proceedings of the 22<sup>nd</sup> International Conference on Software Engineering and Knowledge Engineering (SEKE), 2010, pp. 279–285.
- [6] Jang, W., and G. Shmueli. (2010, July). Modeling online auctions. [Online]. Available: <http://www.modelingonlineauctions.com/datasets>.
- [7] Lei, B., H. Zhang, H. Chen, L. Liu, D. Wang, "A K-means Clustering based Algorithm for Shill Bidding Recognition in Online Auction," in Proceedings of the 24<sup>th</sup> Chinese Control and Decision Conference (CCDC), 2012, pp. 939–943.
- [8] Lin, S.J., Y.Y. Jheng, C.H. Yu, "Combining Ranking Concept and Social Network Analysis to Detect Collusive Groups In Online Auctions," *Expert Systems with Applications*, vol. 39, 2012, pp. 9079–9086.
- [9] Majadi, N., J. Trevathan, and N. Bergmann, "Analysis on Bidding Behaviors for Detecting Shill Bidders in Online Auctions," in Proceedings of International conference on Computer and Information Technology (CIT), 2016, pp. 383–390.
- [10] Mamun, K., and S. Sadaoui, "Combating Shill Bidding in Online Auctions," in Proceedings of International Conference on Information Society (i-Society), 2013, pp. 170–176.
- [11] Patel, R., H. Xu, A. Goel, "Real-Time Trust Management in Agent Based Online Auction Systems," in Proceedings of the 19<sup>th</sup> International Conference on Software Engineering and Knowledge Engineering (SEKE), 2007, pp. 244–250.
- [12] Sadaoui, S., X. Wang, and D. Qi, "A Real-Time Monitoring Framework for Online Auctions Frauds," *Current Approaches in Applied Artificial Intelligence, Lecture Notes in Computer Science*, vol. 9101, 2015, pp. 97–108.
- [13] Sadaoui, S., and X. Wang, "A Dynamic Stage-Based Fraud Monitoring Framework of Multiple Live Auctions," *Applied Intelligence*, vol. 46, no. 1, 2016, pp. 1–17.
- [14] Trevathan, J., and W. Read, "Undesirable and Fraudulent Behavior in Online Auctions," in Proceedings of International Conference on Security and Cryptography, 2006, pp. 450–458.
- [15] Trevathan, J., and W. Read, "Detecting Shill Bidding in Online English Auctions," *Handbook of Research on Social and Organizational Liabilities in Information Security*, IGI Press, USA, 2005, pp. 446–470.
- [16] Trevathan, J., W. Read, "Investigating Shill Bidding Behavior Involving Colluding Bidders," *Journal of Computers*, vol. 2, no. 10, 2007, pp. 63–75.
- [17] Tsang, S., Y.S. Koh, G. Dobbie, S. Alam, "Detecting Online Auction Shilling Frauds using Supervised Learning," *Expert Systems with Applications*, vol. 41, no. 6, 2014, pp. 3027–3040.
- [18] Tsang, S., Y.S. Koh, G. Dobbie, S. Alam, "Span: finding collaborative frauds in online auctions," *Knowledge-Based Systems*, vol. 71, 2014, pp. 389–408.
- [19] Xu, H., and Y. T. Cheng, "Model Checking Bidding Behaviors in Internet Concurrent Auctions," *International Journal of Computer Systems Science & Engineering*, vol. 22, no. 4, 2007, pp. 179–191.
- [20] Xu, H., C. K. Bates, and S. M. Shatz, "Real-Time Model Checking For Shill Detection in Live Online Auctions," in Proceedings of International Conference on Software Engineering Research and Practice (SERP), 2009, pp. 134–140.
- [21] eBay: Number of users 2017. [Online]. Available: <https://www.statista.com/statistics/242235/number-of-ebays-total-active-users/> [Last Accessed on: 16-May-2017]
- [22] Internet Crime Complaint Center. 2014 Internet Crime Report. [Online]. Available: [https://www.fbi.gov/news/news\\_blog/2014-ic3-annual-report](https://www.fbi.gov/news/news_blog/2014-ic3-annual-report), 2015 [Last Accessed on: 22-Dec-2016]
- [23] Internet Crime Complaint Center. 2015 Internet Crime Report. [Online]. Available: [https://pdf.ic3.gov/2015\\_IC3Report.pdf](https://pdf.ic3.gov/2015_IC3Report.pdf), 2016 [Last Accessed on: 22-Dec-2016]